



**University of
Zurich**^{UZH}

**Zurich Open Repository and
Archive**

University of Zurich
University Library
Strickhofstrasse 39
CH-8057 Zurich
www.zora.uzh.ch

Year: 2007

A seeming solution to a seeming puzzle in explaining action

Saporiti, Katia

DOI: <https://doi.org/10.1515/9783110898798>

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-81587>

Book Section

Published Version

Originally published at:

Saporiti, Katia (2007). A seeming solution to a seeming puzzle in explaining action. In: Leist, Anton. Action in Context. Berlin/New York: de Gruyter, 303-314.

DOI: <https://doi.org/10.1515/9783110898798>

A Seeming Solution to a Seeming Puzzle in Explaining Action

Katia Saporiti

1. Hornsby's Claim

Jennifer Hornsby begins her paper "Knowledge in Action" by stating the following truism: "When you learn what someone's reasons were for acting as they did, you know why they acted: you have an explanation of their acting." (285)¹ As is the case with most truisms, this does not tell us much. Think of all the different contexts in which we use the word "why" to raise a question and, again, all the different contexts in which we ask for an explanation. There surely are as many different kinds of things one might wish to know when asking "Why?" as there are different kinds of things one would be inclined to accept as explanations in some context or other. That a reason is something you can refer to in an explanation of something or in an answer to a "why"-question therefore does not say much. What counts as an explanation of an action and what as an answer to the question why a certain person acted the way she did and what counts as a reason for a particular course of action, are open and much debated questions. And in thinking about these questions we should keep in mind that there may be more than one answer to each of them, because there are indeed many different contexts in which we discuss our actions. Not only do we explain actions, we justify and judge them, we endorse them or warn against them, we rationalize them, etc. All these and various other practices we engage in may involve talk about reasons for actions. And what it is that we want to know if in a given context we want to know a reason for an action can always be tested by asking ourselves what kind of answer we would be likely to accept.

Hornsby's main thesis is that acting for a reason essentially involves knowledge on the part of the agent. There is an obvious sense in which this is true: We can attribute beliefs only to subjects whom we credit with knowledge. And only somebody who believes something can possibly act for a reason. But this is, as we will see, not the sense in which Hornsby's thesis is to be understood. Hornsby claims that "we have to allow a role

1 Parenthetical references in the text are to Hornsby 2007 (in this volume).

for knowledge in action if we are to have a correct understanding of what it is to act for a reason" (285). And according to her, it is not knowledge in general that is required on the part of the agent. Rather, she believes that our concept of someone's acting for a reason cannot be explained but with reference to some specific piece of knowledge attributed to the agent.

According to Hornsby, the lack of acknowledgement of the role knowledge plays in explaining action is due to the widely accepted Davidsonian belief-desire-model of reason-explanation. But while philosophers of mind usually take reasons to be pairs of beliefs and desires, moral philosophers tend to take reasons to be facts that favour actions. Hornsby's aim is to reconcile these two different conceptions of reasons. This, she claims, can be achieved only if we take into account the role knowledge plays in action. To her "the things that are agents' reasons are things which agents *know*" (286). Hornsby tries to establish her thesis, first, by introducing a puzzle which she thinks is created by the difference in the conceptions of reasons prevalent in the philosophy of mind, on the one hand, and ethics, on the other, second, by presenting a solution to this puzzle.

In what follows I will argue that the puzzle which Hornsby sometimes calls a seeming puzzle is not all that puzzling. Its construction rests on a confused appeal to the way we talk and (as does the presented solution) the unwarranted claim that different ways of talking about reasons for actions, which are of good use to us when we engage in various practices such as justifying, explaining, excusing or condemning actions and agents, should be unified by an account of what it is to act for a reason. While it is true that such an account must allow for the different ways we refer to reasons and while it certainly is a desideratum to shed light on how these different ways relate to one another, it is not clear why there should be one and only one answer to the question what it is that can be rightly called a reason for an action. (The question itself sometimes seems to be a result of an inappropriate hypostasis.) In the end we will also see that a person who considers the alleged puzzle a real puzzle, in the first place, should not be satisfied with the solution Hornsby presents.

2. The Seeming Puzzle and a First Reaction to It

Hornsby begins to construct her puzzle by observing that "there are plenty of instances" where the following two things hold:

(B) *X* had a reason to ϕ : they believed that *p*.

(F) A reason for *x* to ϕ was that *p*. (286)

Any account of what it means to act for a reason, Hornsby claims, has to endorse this point. And she goes on to point out that while (B) does not seem to imply *p*, (F) in fact does imply *p*. For (F) can only be true if *p* is, while the truth of *p* is not implied by (B)'s being true. This in Hornsby's opinion gives rise to the following puzzle: "It can then seem as if an agent's acting for a reason both did *not* require and *did* require the truth of their beliefs." (287) Later she puts the puzzle somewhat differently:

It could seem then as though when corresponding instances of (B) and (F) are true, an agent has two different reasons for acting [...] a reason which they would have even if (contrary to fact) their belief was false; and another reason which they couldn't have—which would not be a reason—unless their belief was true. (287)

A third formulation of the puzzle offered by Hornsby goes like this: "To *have* a reason, they [the agents] are not required to grasp any truth; but in order for something they have grasped to *be* a reason, it has to be a truth that they've grasped." (287)

A first reaction to this could be the feeling that there is nothing to be puzzled about here. (B) and (F) just are different claims. They can serve to describe different situations. We sometimes act on false beliefs. Those are the cases in which (B) is true, but (F) is not. Luckily, there are plenty of instances where we act on true beliefs. And in these cases both (B) and (F) obtain. In accordance with this first reaction one might try to reformulate (B) and (F) in the following way:

(B') – *X* believed that, given her situation, it would be reasonable to ϕ if *p*.

– *X* believed that *p*.

(F') – *X* believed that, given her situation, it would be reasonable to ϕ if *p*.

– *X* believed that *p*.

– *p*.

Given this reformulation (B) is implied by (F). But (F) is not a consequence of (B). And (F) implies the truth of *p*, while (B) does not. The Davidsonian *dictum* that is at stake here runs as follows: "Actions are events that are caused and rationalized by a desire and beliefs of the agent." In the light of this first reaction to the puzzle one would have to say that what causes and explains the action is the agent's belief that *p* rather than the fact that *p*. What might be felt to be awkward here is that in (F) what is called a reason seems to be the fact that *p*. For one might want to argue that if *p* does not obtain, somebody's acting on the false

belief that p is acting without a reason. Such an agent would be mistaken about having a reason to act the way she does. X thinks she has a reason to ϕ , while in fact she hasn't. The reason in question (namely, that p) does not hold. After all, how can a fact that does not exist (a state of affairs that does not obtain) cause somebody to act in a certain way? But surely, even in the case where it is true that p it is in an obvious sense not the fact that p which causes our agent to act the way she does, but her belief that p —just as it is not the fact that p , but the fact that she believes that p , which explains her action to us (which makes her behaviour a reasonable one). For if it were a fact that p but the agent did not believe, let alone know, that p , we could not say that she ϕ -ed because p .

Does this mean that reasons aren't causes any more? No, it does not mean that. To have a reason to act the way one does just means to act in a way that is caused by and can be explained with reference to one's beliefs and desires. But often, and this is what seems to create the confusion here, in talking about our actions we do not refer to our beliefs and desires but to *what* we believe and desire as reasons we have for acting the way we do. There are different motives for doing this. Sometimes it is just an abbreviated and somehow simplified way to talk, suited to our needs. Often, if an agent acts on a true belief, we do not bother to mention his cognitive states, but only talk about the fact he believes to hold. (We would say "she ϕ -ed because p " rather than "she ϕ -ed because she believed that p ", using the latter kind of expression especially in cases where something went wrong, i.e. the agent acted on a false belief.) And if all goes well, these facts in one way or another are what caused the agent to have the relevant beliefs that made him act the way he did. So reasons are causes, all right, either in the sense that they prompt the action, if we talk about beliefs and desires, or in the sense that they are at the origin of a causal chain that led the agent to have the beliefs and desires that prompted the action. (Obviously, there are very different concepts of causation involved here.)

In other contexts, though, we are not so much concerned with the agent's beliefs (nor are we much concerned with what made him act the way he did) but with looking at an action in the light of the facts. This is often the case when we want to judge an action (or an agent for what he did or did not do). We often discuss different courses of action, talking about their pros and cons, thereby in a sense abstracting from the individual agent. Also, if somebody took a particular course of action despite there being good reasons not to do so, or if somebody didn't do what there were good reasons to do, we will ask questions as to whether

he knew of these reasons or whether he could have known better etc. Our interests in these questions are manifold. Sometimes we want to judge the agent, at other times we want to find out how to improve a kind of behaviour in the light of certain goals etc. But what we refer to as a reason in this sort of discussion cannot straightforwardly be regarded as something which causes an action. Very often the actions whose reasons we are talking about have not even taken place. This does not mean, of course, that our talk of reasons here does not have anything to do with the reasons which explain and cause behaviour. Quite to the contrary, the different practices we engage in when we talk about actions and reasons for actions are involved and mutually dependent. Roughly speaking, we could not explain someone's behaviour by referring to his belief that p unless we could regard p as favouring his action. And we could not regard a state of affairs as favouring an action, unless *ceteris paribus* we ourselves were inclined to take the respective action in case we believed the state of affairs to obtain.

3. Resolving the Puzzle

Setting aside our first reaction to Hornsby's puzzle, one may want to modify the reconstructions of (B) and (F) that I have given. (B), "X had a reason to ϕ : they believed that p ", could again be spelled out as:

- (B') – X believed that, given her situation, it would be reasonable to ϕ if p .
- X believed that p .

But (F), "A reason for x to ϕ was that p ", would come out as something like

- (F*) – For anybody in the same situation as x it would be reasonable to ϕ if p .
- p .

(F*) would amount to the assertion that there was good reason for x to ϕ , whether she knew it or not, and whether she in fact did ϕ or not. (F*) does not entail that x ϕ -ed. Nor does it entail that x believed that p . It does not even entail that, had x believed or known that p , she would have ϕ -ed. (F*) also does not entail that x believes that it would be reasonable for anybody in her situation to ϕ if p . In fact (F*) does not seem to tell us much about x at all. Instead, it tells us something about the situation we take x to be in and about what we would assume to be a reasonable reaction to p or what, given the situation of x , would count as rational

behaviour. But in asserting something along the lines of (F*) we do treat x as a rational agent; we assume that, had she believed that p , she *ceteris paribus* would have ϕ -ed. We assume that, had she believed that p and indeed ϕ -ed, we would have been able to explain her behaviour by saying “she ϕ -ed because p ”. We treat x as being one of us, as engaging in our practice of explaining, evaluating, judging and predicting actions. If (F*) were true, then, if x believed that p and didn’t ϕ , we would be entitled to putting forward a hypothesis like “ X didn’t, after all, believe p to be a reason to ϕ (and there she was mistaken)” or “ x had more powerful reasons not to ϕ ”.

Note that (B’) and (F’), which resulted from the first reaction to the puzzle, stand in different relations from (B’) and (F*). Again, it could easily be the case that (B’) holds while (F*) does not. But also one can easily imagine cases where (F*) holds while (B’) does not. What is likely to be the most important point, according to Hornsby, is that there are cases, probably ideal cases, we think of when we give belief-desire-explanations of our actions, where both (B’) and (F*) hold.

(F*), which I suggested as a reformulation of what could be meant by (F), certainly is not a reconstruction Hornsby had in mind, as she assumes throughout that from (F) we can derive an explanation of the form “ x ϕ -ed because p ”. But this surely presupposes not only that p and that x did indeed ϕ , but also, among other things, both that x believed that p and that x believed that, given her situation, it would be a good idea to ϕ if p . Furthermore, it presupposes that x ϕ -ed on these beliefs and not for some other reason. But we are not concerned here with giving a full, let alone a reductive, account of what it means to act for a reason. So we can set aside many of these complications. In concentrating on the explanation of actions that have taken place, Hornsby tends to overlook that the way we postulate the existence of reasons for actions nobody ever performs is instructive as regards understanding explanations along the lines of (F).

Having started off with the widely shared assumption that “ x has a reason to ϕ ” means that “ x ’s ϕ -ing is caused by and can be explained by x ’s beliefs and desires”, we can now see how we might want to adjust this. As (F*) implies neither that x ϕ -s nor that she has the belief that p , we might want to say that “ x ’s having a reason to ϕ ” means that x ’s ϕ -ing, if indeed she did ϕ in a given situation, could be explained and caused by her beliefs and desires. And we could say that “ p ’s being a reason for x to ϕ ” means that x ’s ϕ -ing, if indeed she did ϕ in a given situation, could be explained and caused by her belief that p , if indeed she believed that p .

What about the puzzle Jennifer Hornsby has drawn our attention to? (F*) does imply the truth of p , while (B') does not. Does this mean that having a reason to ϕ at the same time requires having a true belief one acts upon and does not require having a true belief one acts upon? Or, to raise a slightly different, but clearly related, question: Do we use the word "reason" in different ways, does it have different meanings in (B) and (F)? And if so, does this undermine our understanding of reason-explanation? Hornsby insists that any account of what it means to act for a reason has to allow for the fact that there are plenty of cases where both (B) and (F) yield true sentences. This certainly is true. But throughout her paper she also seems to assume that this means that we have to come up with an account that not only allows (B) and (F) to yield true sentences but which in addition to that gives us one core-meaning of "reason" and a corresponding canonical use of the term that could be represented by another schema which would be entailed by both (B) and (F). This canonical use would lie at the bottom of all other uses of the term, which we would then have to regard as derived from it. If by using "reason" in (B) and (F) we refer to different things—e.g. beliefs and facts—'reason-explanation' becomes unintelligible, or so Hornsby seems to think. But this worry is unfounded. While any account of reason-explanation should allow for a variety of well-established practices and, if possible, explain their relations, it certainly does not have to single out one sort of entity which we are justified in calling a reason. Neither is there a reason for assuming that from all the different things we do when we talk about actions and the reasons agents can have or should have for performing them, one can be singled out as underlying all the others.

4. Constraints on an Account of Reason-Explanation

In presenting her puzzle and dismissing several solutions to it Hornsby tries to spell out some of the constraints an account of reason-explanation has to meet. A discussion of Ryle's work leads her to maintain that (B) and (F) should not be treated as if they yielded different kinds of explanations and that an agent's knowing that p should secure his believing that p . While the latter seems uncontroversial, the former claim is somewhat unspecific. If what is meant is that (B) and (F) should not be treated as excluding each other in the sense that they cannot be true at the same time, then it certainly is true. If what is meant is that there is no difference between (B) and (F), the claim is certainly false, as (F) has

implications which (B) does not have, namely the truth of p . This, after all, is how the puzzle is supposed to arise.

Hornsby comments on Jonathan Dancy's observation that it is perfectly all right to say things like "His reason for doing it was that it would increase his pension, but he was sadly mistaken about that". From this Dancy, according to Hornsby, wants to draw the conclusion that reason-explanations as in (F) are non-factive. Hornsby objects that this does not take into account "that when a reason for doing something is that p , it is true that p " (291). But while Dancy may not succeed in showing that all reason-explanation is non-factive, his observation that we do say things along the following line, is certainly correct: " X 's reason to ϕ was that p , but he was mistaken about p ".

This reveals something Hornsby tends to overlook. When we refer to a certain state of affairs as a reason for an action, we sometimes imply that this state of affairs obtains and sometimes we don't. And this is not, as Hornsby assumes, tied neatly to certain ways of expressing ourselves. Sometimes it is possible to say that p *was* the reason for x to ϕ , or that x ϕ -ed because of p etc., even though it was not the case that p . Therefore it is not a 'datum' "that when a reason for doing something is that p , it is true that p ", as Hornsby claims. Equally, we may in certain situations say that p is a reason *for* x to ϕ and accordingly that x *has* a reason to ϕ , even if x does not believe or know that p . Speech is not as tidy as Hornsby supposes when she confines herself to using the expression "to have a reason" in talking of beliefs that may be true or false, and to using the expression "to be a reason" in talking of facts that favour actions.

Hornsby accuses Dancy of inappropriately assimilating an agent who acts on a true belief to an agent who acts on a false belief. She claims that "what we want to acknowledge is not only that people may have reasons for acting whether they believe truly or falsely, but also that people sometimes act for reasons there really are." (292) But Hornsby does not succeed in explaining why this cannot be done by simply adding that the belief we refer to in explaining somebody's action is true. Also it remains unclear why we should not want to assimilate somebody's acting on a true belief to somebody's acting on a false belief. Besides, talking of "reasons there really are"—presumably distinguishing them from reasons which are not really there—is an illegitimate hypostasis if it leads us into thinking of entities which are justly called reasons and which exist independently of a practice of invoking them. Hornsby puts her point somewhat differently when she asks: "Do we not want to say that, when all is well, that in the light of which one acts must be among the reasons

in favour of so acting?” (292, Fn 5) But here again she supposes that only a fact can favour an action. Why should we not allow other things to favour our actions (states of affairs which do not obtain or beliefs, desires, wishes, moods, habits, emotions, aims, goals, norms, laws, rules, challenges, provocations, etc.)?

One reason for thinking this may be brought to light by looking at Hornsby's discussion of Michael Smith's distinction between normative and motivating reasons and Stout's claim that reason-explanations are not causal explanations. Hornsby worries that if we separate whatever is referred to as a reason in (B) from what is referred to as a reason in (F) we render reason-explanation unintelligible in that what we get in (B) is the agent's motivation while what we get in (F) is the rationale of the action. So it seems that only (B) and (F) together can reveal an agent as being rational. If we disconnect reasons in (B) from reasons in (F), we would either have to give up on the motivational aspect or on the rational aspect of reasons, or so Hornsby thinks. This way we would not be able to ever regard a person as being under the influence of reason. From her discussion of Smith and Stout it becomes clear that Hornsby assumes that to be under the influence of reason an agent has to be influenced by a reason, where this means that for an agent to act rationally there has to be a fact that favours and causes his action in a straightforward way. But surely, a person who acts on a false belief can be perfectly rational. And by assuming that we can state diverse things by (B) and (F) and that there are many different things that can be called a reason, we do not commit ourselves to the view that (B) and (F) are completely disparate and that there is nothing that connects the various things we can count as reasons.

To summarize: We have to allow that there are many cases where (B) and (F) hold without giving rise to Hornsby's puzzle. It seems that we can avoid doing so only if we assume that (B) and (F) yield different explanations and that what is called a reason in (B) is not the same as what is called a reason in (F). At the same time, though, we cannot allow explanations along the lines of (B) and (F) to exclude each other. And although we want to make a difference between somebody who acts on a false belief and somebody who acts on a true belief, we cannot allow (B) and (F) to drift apart in a way that renders it unintelligible that people who act for reasons are rational in what they do. Instead, we have to account for the fact that (B) and (F) are indeed closely related. Furthermore, and this is something I do not have room to discuss but which seems to me a convincing claim Hornsby makes in discussing

Stout: we must not assume that for an agent to act for a reason he needs to have mastered the concept of a reason.

5. The Seeming Solution

Let us now look at how Hornsby tries to accommodate her findings by presenting a solution to her puzzle. She draws our attention to the connection she believes to hold between p 's being an agent's reason to act the way he does and the agent's knowing that p . She maintains that, unless x knows that p , p cannot be his reason for ϕ -ing. (On the other hand, she claims that p could be a reason for him to ϕ (296).) Somebody who has grasped a reason has knowledge, or so she maintains. This, according to Hornsby, is one of the things we have to bear in mind if we want to solve the puzzle. She construes a Gettier-like example in which an agent's true and justified belief that p does not amount to knowledge, and she claims that in this case where p is indeed a reason for x to ϕ (i.e. p is a fact that favours his ϕ -ing) and in which x indeed ϕ -s because he believes that p , we still cannot record x 's reason to ϕ by saying " x ϕ -ed because p " (296). Hornsby does not explain, though, why exactly we should not allow this last formulation to be a correct description of the situation. She just compares the agent to somebody who acts on a true and justified belief which undoubtedly amounts to knowledge as everything is normal and not Gettier-like. Hornsby claims that "knowledge is what makes the difference here. This is because the difference that needs to be made is just the sort of difference that knowledge makes. The factive character of knowledge ensures that a person's state of mind, if it is knowledge, relates them to actual reasons." (297) There are a number of obvious responses to this. 1) Gettier-like situations are exceptions and their bearing on a general practice is doubtful. 2) One may doubt that Gettier-examples prove that knowledge involves more than justified true belief, as they describe situations where it may be doubted that the relevant belief is truly justified. 3) It remains unclear what it could be that knowledge adds so that we are allowed to say " x ϕ -ed because p " only if x knows that p . Surely, a true belief is as factive as knowledge!

The second point we have to bear in mind to solve the puzzle is, according to Hornsby, that somebody who believes that p treats p as if he knew that p . Considering the following schemata of explanations

- (b) x ϕ -ed because they believed that p .
- (k) x ϕ -ed because they knew that p .
- (f) x ϕ -ed because p . (299)

Hornsby suggests to rewrite (b) as “ X ϕ -ed because they treated p as they would if they knew that p .” By doing this we are supposed to arrive at a better understanding of the way somebody’s acting on a belief is related to somebody’s doing something for which what he believes *is* a reason. “By introducing (k), we make a connection between the notion of ‘reason’ as it occurs in (B) and in (F), and we see why there should be a single reason at work when instances of both obtain.” (300) What we still do not see, though, is what is accomplished by introducing knowledge that cannot be accomplished by acknowledging that (F) and (f) can only be true if the agent’s belief that p is true. Hornsby solves her puzzle by ensuring that the truth of (F) implies the truth of (B) and correspondingly the truth of (f) implies the truth of (b). But this, as we have seen earlier, can easily be done by plausible if not complete explanations along the lines of (B’) and (F’). And anybody who believed in ‘double-counting reasons’—that is, believed that in instances where (B) and (F) both yield true sentences we are dealing with two different reasons the agent had for acting—will still believe so. He would try and argue just as Hornsby has done that x has one reason (namely his believing that p or his treating p as if he knew that it were true) which he would have even if it were false that p , and another one which he would not have if it were false that p (namely p —the fact which favours his action).

On the face of it, the assumption that the difference (if there is one) between p ’s being a reason for x to ϕ and p ’s being x ’s reason to ϕ lies in x ’s knowing that p does not seem plausible at all. If for p to be a reason for x to ϕ x does not have to know that p , then neither does x need to know that p for p to be x ’s reason to ϕ . X just has to believe that p and p has to be true. But a true belief does not amount to knowledge. Furthermore, believing that p does not amount to treating p as something one knows; nor is believing that p to the believer in some sense indistinguishable from knowing that p . If this were the case, why should anybody ever try to find out which of her beliefs are true? What is true, though, is that believing that p just is taking it to be true that p . This corroborates the suspicion that regarding the alleged puzzle there is nothing to be gained from invoking knowledge.

